

6 - Questions-Réponses

? Vous donnez votre avis et vous justifiez . = Vous inscrivez ce qui est attendu

a?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

Les deux séries donnent le même résultat.

```
A1 > plot(x,y)
A2 > points(x,x-2,pch=2)

B1 > points(x,x-2,pch=2)
B2 > plot(x,y)
```

b?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

C'est impossible.

```
> rnorm(1)
[1] 0.7741
> 2*rnorm(1)
[1] 0.6702
```

c?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

Il y en a au moins un de ces ordres qui va générer une erreur.

```
> glm(w~x,family="binomial")
> glm(w~x,family="gaussian")
> glm(w~x,family=gaussian)
> glm(w~x,family=binomial)
> glm(w~x)
> glm(w~x,family=poisson)
```

d=

```
> x<-seq(9,1,-1)
> x
[1] 9 8 7 6 5 4 3 2 1
> x<-matrix(x,nrow=3)
> x
      [,1] [,2] [,3]
[1,]    9    6    3
[2,]    8    5    2
[3,]    7    4    1
> x[rank(x[,1]),]
```

e=

```
> x<-1:5
> y<-rep(c(0,1),c(2,3))
> plto(x,y)
```

f?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

C'est vraisemblable.

```
> x<-rnorm(4)
> x
[1] 0.3173 -2.4535 1.4957 -5.4490
```

g?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

Ceci est invraisemblable.

```
> r<-rbinom(10,1,0.05)
> r
[1] 1 1 1 1 1 1 0 1 1 1
```

h=

Lequel des deux ordres est valide A ou B ?.

```
> is.data.frame(donax)
[1] T
> donax[1,] # ordre A
> donax(1,) # ordre B
```

i=

```
> y<-rep(c(0,1),c(2,3))
> y
```

i=

```
> quantile(x,probs=seq(0,1,0.1))
 0%  10%  20%  30%  40%  50%  60%  70%  80%  90% 100%
 1  1.9  2.8  3.7  4.6  5.5  6.4  7.3  8.2  9.1  10
> quantile(x,probs=0.8)
```

k=

```
> range(100:101)
```

l=

```
> x<-rpois(100,3.25)
> max((x-min(x))/(max(x)-min(x)))
```

m?

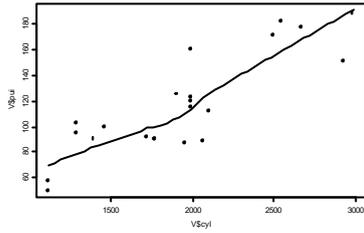
• *c'est vrai* • *c'est faux* • *je n'en sais rien*

la figure A est générée par :

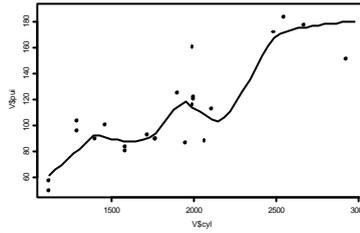
```
> scatter.smooth(V$cyl,V$pui,span=0.4)
```

la figure B est générée par :

```
> scatter.smooth(V$cyl,V$pui,span=0.8)
```



A



B

n?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

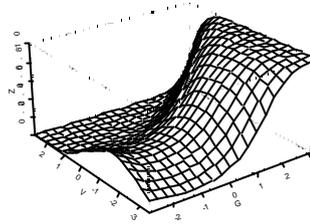
cauchy est fonction de S-PLUS qui permet d'utiliser les distributions de Cauchy.

o?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

Cette figure est obtenue à l'aide d'une des fonctions suivantes :

lm, contour, prespp, points, plot, draw, graph3d, freqgrad, predict.



p=

> 2^2*3^2

q=

Quel nombre entier est utilisé dans ce dialogue ?

```
> x<-e+rnorm(200) # e est un entier
> mean(x)
[1] 5.045
> var(x)
[1] 0.9867
```

r?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

La variance de \mathbf{x} sur n observation est la moyenne des n variances de \mathbf{x} calculées sur les n sous-groupes à $n - 1$ observations.

```
> var1<-function(x){return(var(x,unbiased=T))}
> var2<-function(x){return(var(x,unbiased=F))}
> jackknife(V$cyl,var1)
Call:
jackknife(data = V$cyl, statistic = var1)
```

Number of Replications: 24

Summary Statistics:

	Observed	Bias	Mean	SE
var1	278688	0	278688	74739

```
> jackknife(V$cyl,var2)
```

Call:

```
jackknife(data = V$cyl, statistic = var2)
```

Number of Replications: 24

Summary Statistics:

	Observed	Bias	Mean	SE
var2	267076	-11612	266571	71490

S=

Qu'est-ce qu'il attend pour rendre la main ?

```
> plot(0,0)
> legend(locator(1), "coucou")
```

t?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

Si lm1 est un modèle linéaire, il y a plusieurs manières de rendre utile l'ordre

```
> plot(lm1)
```

u?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

La probabilité de dépasser 2.776 pour une loi t à 4 degrés de liberté est de 5%.

V=

```
> 2*T
[1] 2
> 2*F
[1] 0
> x<-c(T,T,F,F,T)
> y<-1:5
> x*y
```

W=

Comme dans le précédent :

```
> is.numeric(x)
> is.logical(x)
```

X=

```
> sort(sample(c(1,2,3,4,5),5,replace=F))
```

V=

```
> 3*length(cumsum(sample(10,10)))-5
```

Z?

• *c'est vrai* • *c'est faux* • *je n'en sais rien*

La fonction la plus utile de S-PLUS est :

```
> quit()
```

Q 1 Que vous inspire l'affichage suivant ?

```
> 2*rnorm(1)
[1] -6.4829
```

Q 2 Quel est le résultat ?

```
> print(pi,digits=12)
```

Q 3 L'ordre ci-dessous peut-il renvoyer la valeur 0 ?

```
> sum(rbinom(900,1,0.01))-rbinom(1,900,0.01)
```

Q 4 A quelle condition le résultat suivant est-il possible ?

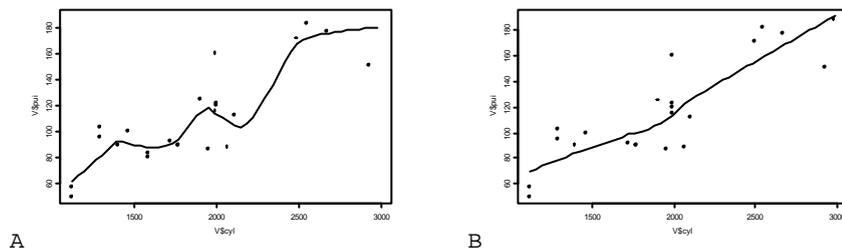
```
> sort(sample(c("a", "b", 12.7, "c", n0, "aaa", "bb"), 6, replace=F))
[1] "10" "a" "aaa" "b" "bb" "c"
```

Q 5 Quelle valeur a t'elle été affectée à n0 ?

```
> 3*length(cumsum(sample(n0)))-5
[1] 7
```

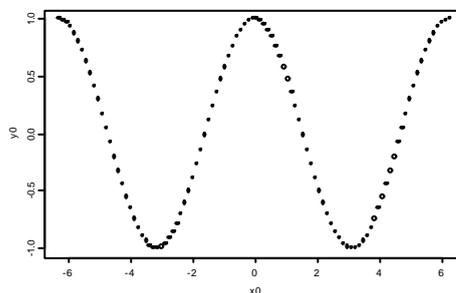
Q 6 A-t-on nécessairement $0 \leq a < b \leq 1$ pour obtenir ce qui suit ?

```
> scatter.smooth(V$cyl, V$hwy, span=a) # résultat A
> scatter.smooth(V$cyl, V$hwy, span=b) # résultat B
```



Q7 Quelle chaîne de caractères a-t-elle été utilisée pour xxxxx ?

```
> x0<-seq(-2*pi,2*pi,le=100)
> y0<-xxxxx(x0)
> plot(x0,y0)
```



Q8 A quel résultat vous attendez vous ?

```
> quantile(x, probs=seq(0,1,0.1))
 0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
 1 1.9 2.8 3.7 4.6 5.5 6.4 7.3 8.2 9.1 10
> (quantile(x, probs=0.73)) < 8.2
```

Sur 38 points d'écoute ornithologique on a extrait les données sur deux espèces, la pie et la corneille. Les résultats sont :

	Corneille	
	absente	présente
Pie absente	13	8
Pie présente	6	11

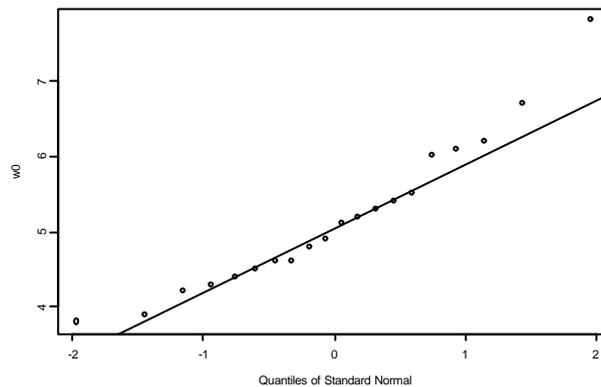
Q9 Quelle est la valeur du Chi2 de cette table de contingence et sa signification ?

Lors d'une étude de croissance, la hauteur de 20 plantules a été mesurée (en cm) :

4.6	7.8	4.8	6	5.4	3.9	4.5	6.2	4.9	5.3	6.7	4.2	3.8	5.1	6.1	5.2	4.3	4.6	4.4	5.5
-----	-----	-----	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Q10 Peut-on considérer que cette variable suit une loi normale ? (Note : voir ks.gof)

Q11 Quels ordres ont-ils donné ce graphique ?



Dans une conversation autour d'une méthode sur l'analyse des niches écologiques M. Montadert propose un tableau très caractéristique du problème des courbes de réponse aux variables environnementales. Dans $n = 102$ stations on connaît la présence ou l'absence (0/1) de la gélinotte (geli) et un descriptif de la végétation avec (entre autres) les variables abondance du Framboisier (fra), abondance du Noisetier (noi) et abondance du Sorbier (sor). fra, noi et sor sont les variables explicatives. geli (binaire) est la variable à expliquer.

Pour contrôler l'implantation des données, utiliser le tableau 1. Le fichier G.txt est disponible dans le répertoire indiqué oralement et sera lu par un ordre du type :

```
> G<-read.table("--- G.txt",h=T,sep=",")
> G
      geli fra noi sor
1      0 0.75 3.00 0.75
2      0 1.00 2.75 3.00
3      0 1.00 2.75 1.25
...
100    1 1.00 1.50 2.00
```

101 1 0.50 0.75 2.00
 102 1 0.50 0.75 3.25

Q12 Donner les moyennes des variables fra, noi et sor.

Q13 La corrélation entre l'abondance du Sorbier et celle du Noisetier est-elle significativement non nulle ?

Q14 Quelle est la première valeur propre de l'analyse en composantes principales normée du tableau des trois variables d'habitat ?

On appelle tot la variable qui attribue à une station la somme des trois explicatives :

```
> tot<-G$fra+G$noi+G$sor
```

La corrélation entre la première coordonnée de cette ACP et la variable tot montre que la somme simple suffit largement à synthétiser les trois explicatives.

Q15 Quelle est cette corrélation ?.

	fra	noi	sor
1	1.25	3	0.25
2	1.25	3	3
3	0.75	2.5	2.75
4	1	1.5	4
5	0.25	2.5	2.5
6	0.75	1.75	2
7	1.25	1	3.25
8	1	2.5	2.5
9	1.25	2.25	3.75
10	0.5	2	3.5
11	0	2	1.25
12	0.5	3	2.75
13	0.5	1.5	2
14	1	2.25	2.75
15	0.5	3	2.25
16	0	2.33	2.33
17	0.75	1.25	1.25
18	0.5	0.75	1.25
19	1	1.5	2
20	0.5	0.75	2
21	0.5	0.75	3.25

	fra	noi	sor		fra	noi	sor		fra	noi	sor
1	0.75	3	0.75	28	0	1	1.5	55	1	1.25	2.5
2	1	2.75	3	29	0	0.25	1	56	0.25	0.75	1.5
3	1	2.75	1.25	30	0	1.67	1.67	57	0.5	1.5	2.75
4	1	3	3.25	31	0	1.25	1.25	58	0	2	2.75
5	0	1.25	0.25	32	0.25	3	2.75	59	0	1.5	2
6	1	1.25	1.75	33	0.25	0.25	0.75	60	0	0	0.25
7	0.25	1.75	3	34	0.5	0.75	0.75	61	0.25	0.25	0.75
8	0	1	0.75	35	0.75	0	0.75	62	0.25	0	1.5
9	1	1	0.5	36	1	0	1.25	63	0.25	0.5	1.75
10	1	3	1	37	1	1.25	3.25	64	0.5	0.75	1.75
11	0.5	2.25	2.25	38	0.25	2.25	2	65	0.75	1	1.25
12	0.75	0.75	2.75	39	0	1	1.5	66	1	1	1.5
13	1	0.5	0.5	40	0	2.25	2.25	67	0.5	0	1
14	0	1.75	1.5	41	0	0.75	1.5	68	0.5	1	1.5
15	0.5	1.25	1.75	42	0	0.75	2	69	0.5	0.25	1.75
16	0.5	1.25	2.25	43	0.25	1	1	70	0.25	0.5	1.25
17	0.75	1.75	3.75	44	0	1.33	2.33	71	0	1	2.75
18	1	1.5	3	45	0	1.25	2.5	72	0.25	0.75	1.75
19	1	1.75	3	46	0.25	0	1.25	73	0.25	0.5	1.5
20	1.25	2	2.5	47	0	0.25	3.25	74	1	1	0.75
21	0.25	0	1.25	48	1.5	2	2.75	75	1	0.5	2
22	1	2.25	2.5	49	0.5	0.75	3.25	76	0.75	0.75	3.25
23	0.25	0	0.75	50	1	1	2	77	0.25	0.75	2.25
24	0	2.75	3	51	0.5	0	1	78	0.5	1.25	3
25	0.25	2	2.5	52	0.75	0.25	0.25	79	0	0.25	0.75
26	0.25	1.25	2.75	53	0.75	0.25	1.5	80	0.75	2.75	1.5
27	0.25	0.5	1.5	54	1	1	1.75	81	0.67	3	1.33

Tableau 1 : A gauche 21 stations avec Gelinotte, à droite 81 stations sans Gelinotte.

Q16 Quelle est la fréquence de la présence de la Gélinothe pour les stations où tot est inférieure (respectivement supérieure ou égale) à 4 ?

Q17 La régression logistique montre-t-elle que la probabilité de rencontrer la Gélinothe dépend de chacune des explicatives ?

On a alors deux possibilités :

```
> tot<-G$fra+G$noi+G$sor  
glm(G$geli ~ tot, family = binomial)  
glm(G$geli ~ G$fra+G$noi+G$sor, family = binomial)
```

Q18 Avez-vous une préférence ?

La suite s'en tiendra au modèle

```
> glm0 <- glm(G$geli ~ tot, family = binomial)
```

La probabilité estimée de rencontrer la gélinothe pour une station où la note cumulée des trois explicatives vaut 0 est donnée par :

```
> predict(glm0,newdata=data.frame(tot=0),type="response")  
1  
0.01312
```

Les paramètres du modèle sont :

```
> glm0$coefficients  
(Intercept) tot  
-4.32 0.6913
```

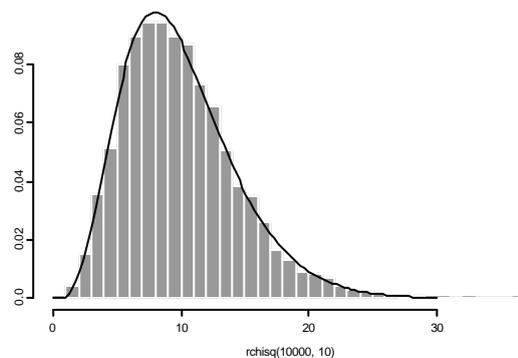
Q19 Quelle relation existe-t-il entre les deux ?

Q20 Reporter sur l'emplacement prévu à cet effet la courbe de réponse de la Gélinothe à la variable tot.

Pour les amateurs :

Q21 Donner le F de l'analyse de variance de la variable tot pour le facteur présence-absence de la gélinothe.

Q22 Quelle est la loi de probabilité représentées ci-dessous ?



Q23 Comment a été tracée la courbe ?